

## **3.9 Science Data Ingest Activities**

### **3.9.1 TRMM Level 0 Data Ingest Scenario**

The following scenario discusses the operations staff involvement in the ingest of TRMM Level 0 data from the Sensor Data Processing Facility (SDPF). This scenario will be implemented at the LaRC and MSFC DAACs.

The Data Ingest Technician is the primary operations staff member involved in the automated Level 0 data ingest process.

#### **3.9.1.1 Scenario Description**

The following scenario is assumed to occur during a given day of the TRMM Mission (Release A) period at the LaRC and MSFC DAACs. Three subscenarios are to be discussed:

- Data Ingest Technician monitoring of ongoing ingest request processing (including the Level 0 ingest request processing)
- Data Ingest Technician cancellation of an ingest request
- Data Ingest Technician viewing of summary ingest request processing status and statistics

The LaRC and MSFC DAACs expect to receive one delivery of TRMM Level 0 data per day. The size of the delivery to both sites is approximately 250 megabytes. An additional smaller data volume (approximately TBD MB) is delivered up to three times per day for use as expedited data. In both cases the delivery is accomplished by means of the EOSDIS Backbone network (EBnet).

Note: in the event of a long-term failure of the network interface between the SDPF and ECS, an 8mm tape backup mechanism is available. The hard media ingest scenario (3.9.4) discusses operations staff procedures for that contingency scenario.

#### **3.9.1.2 Frequency**

In the nominal case the deliveries are expected at regular times during the staffed portion of the day. However, the system is set up to accept deliveries at any time of the day, including unstaffed periods. The LaRC and MSFC DAACs expect to receive one delivery of TRMM Level 0 data per day and up to three deliveries per day of expedited data. Due to the low data volume, the time to process a given delivery of Level 0 data is on the order of minutes in the TRMM Mission (Release A) era.

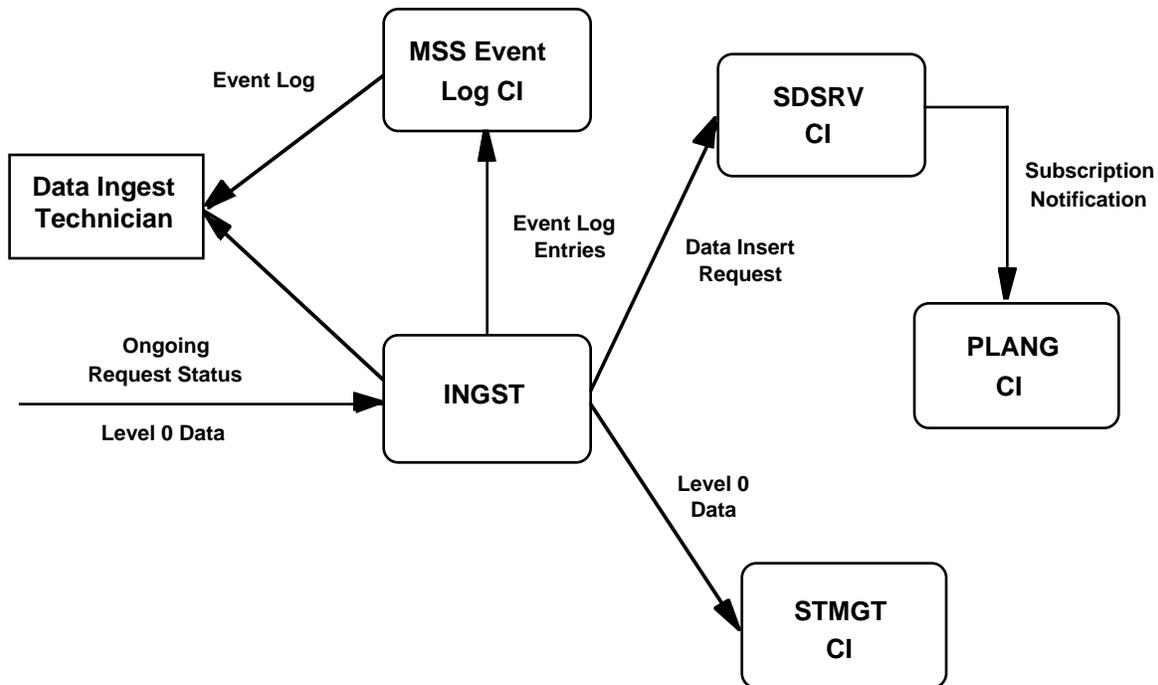
#### **3.9.1.3 Assumptions**

Assumptions underlying this scenario are as follows.

1. The system at the DAAC is in stable operations.
2. The nominal TRMM Level 0 data ingest process is discussed in this scenario. "Fault" conditions are discussed in the TRMM data ingest fault scenario (3.9.2).

### 3.9.1.4 Components

There are five configuration items (CIs) involved with this scenario. Figure 3.9.1.4-1 indicates the interaction between the operations personnel, the INGST CI, the TBD MSS Event Log CI, the SDSRV CI, the STMGT CI, and the PLANG CI.



**Figure 3.9.1.4-1. TRMM L0 Ingest CI/Operations Staff Interaction**

### 3.9.1.5 Preconditions

The following preconditions are assumed for this scenario:

1. Data and metadata configuration information has been documented in an Interface Control Document (ICD) or other operational agreement with the data provider and implemented in advance in ECS tables to appropriately describe the structure of each type of incoming Level 0 data and the processing to be performed on the data.
2. All necessary technical and management personnel have been trained in the use of MSS event log and INGST status monitoring tools. Required skills are outlined in DID607.

### 3.9.1.6 Detailed Steps of Process

Table 3.9.1.6-1 represents the details of the Level 0 data ingest activities. The time scales indicated are approximate. The "User" in this case represents the external data provider, the SDPF. The "Operator" is the Data Ingest Technician, who is responsible for monitoring ongoing data ingest processing.

**Table 3.9.1.6-1. TRMM Level 0 Data Ingest Scenario Process (1 of 3)**

<b>Step</b>	<b>Time*</b>	<b>User</b>	<b>Operator</b>	<b>System</b>	<b>Associated Figure</b>
0	Prior to start of TRMM mission		The TBD operations staff enters a subscription requesting notification upon receipt of TRMM Level 0 data.	The System (SDSRV CI) stores the subscription pending receipt of TRMM Level 0 data.	Subscription setup GUI TBD.
1	Start of work day (2 min)		The Data Ingest Technician invokes the Ingest Status Monitor tool (GUI display) from the main Ingest GUI screen.	The System identifies ongoing ingest requests (stored in Sybase tables) and displays them.	3.9.1.6-1, 3.9.1.6-2
2	L0 data available (5 min)	SDPF application software automatically sets up to transfer Level 0 data to ECS. A Data Availability Notice (DAN) is sent to ECS. A Data Availability Acknowledgment (DAA) is received.		The System automatically checks the received DAN and returns a DAA. Once a request to ingest Level 0 data is received, the INGST CSCI automatically checkpoints request information extracted from the DAN into a Sybase data base. The System automatically coordinates the Level 0 data transfer with the SDPF using ftp services.	
3	L0 data available + a few minutes (10 min)		The Data Ingest Technician periodically reviews Ingest Status Monitor display. The Technician looks for ingest requests that have been queued for an unexpected period (in the TRMM Mission era, anything on the queue for more than a few minutes).	The System automatically extracts metadata from transferred Level 0 data, checks the metadata (e.g., range checks), and inserts the data and metadata into the Level 0 Data Server component of the Ingest Subsystem. Request state (active, file transferred, data insertion complete, etc.) is updated in the checkpointed request information.	3.9.1.6-2

<b>Step</b>	<b>Time*</b>	<b>User</b>	<b>Operator</b>	<b>System</b>	<b>Associated Figure</b>
4	(2 min)		The Data Ingest Technician may cancel a request if desired by selecting Control from the main Ingest GUI screen, selecting the request to be canceled, and selecting the cancel option from the control screen.	The System accepts cancellation request based on a given request ID.	3.9.1.6-3
5	Once every 15 min		The Data Ingest Technician periodically reviews the MSS Event Log to visually determine anomalous conditions (e.g., a pattern of metadata check errors).	The System automatically logs events by means of the MSS Event Logging capability. "Events" include detection of out-of-range metadata values, incompletely-transferred data files, etc. Based on DAAC policy, selected events may be identified as "alerts", which trigger a visual change of state at the MSS Event Log display.	MSS Event Log displays TBD
6	(5 min)	SDPF application software receives a Data Delivery Notice (DDN), indicating the status of the Level 0 data ingest. The SDPF application software transmits a Data Delivery Acknowledgement (DDA) to acknowledge receipt of the DDN.	The Data Ingest Technician observes the removal of the completed ingest request from the Status Monitor display.	Upon completion (successful or unsuccessful) of data insertion into the Data Server, status is automatically returned to the data provider (SDPF) by means of the DDN. After receipt of a DDA, on-going ingest request information is deleted. Summary information is retained in the Sybase data base (as Ingest History Log data).	

Step	Time*	User	Operator	System	Associated Figure
7	(5 min)			Upon completion successful of data insertion into the Data Server, the SDSRV CI automatically determines the existence of subscriptions on the receipt of the TRMM Level 0 data. A subscription notice is sent to the requesting entity.	
8	After ingest request completion (1 min to view log, 5 min to generate report).		The Data Ingest Technician may view summary information about completed ingest requests using the GUI Ingest History Log tool.	The System provides access to Sybase data base tables containing summary information on completed ingest requests, including completion status, data volume ingested, etc.	3.9.1.6-4
9	After ingest request completion (5 min).		The Data Ingest Technician may generate a summary report on completed ingest requests. The report (in two parts) gives summary statistics (e.g., number of data granules ingested, data volume ingested) and error statistics (e.g., number of errors of a given type encountered) for a specified time range.	The System provides access to Sybase data base tables containing summary information on completed ingest requests, including completion status, data volume ingested, etc. Report type and frequency will be based on DAAC policy.	3.9.1.6-5 and 3.9.1.6-6

\*Nominal time required for completion of the selected event is included in braces (minutes).

### 3.9.1.7 Postconditions

At the completion of the above scenario, the Planning Subsystem (PLANG) is notified of availability of the data once all Level 0 data and associated ancillary data are available. The Processing Subsystem (PRONG) is then invoked by the Planning Subsystem (PLANG) for all data against which a subscription has been placed. PRONG will retrieve data from the Level 0 Data Server and process to higher levels per scenario 3.13.1.



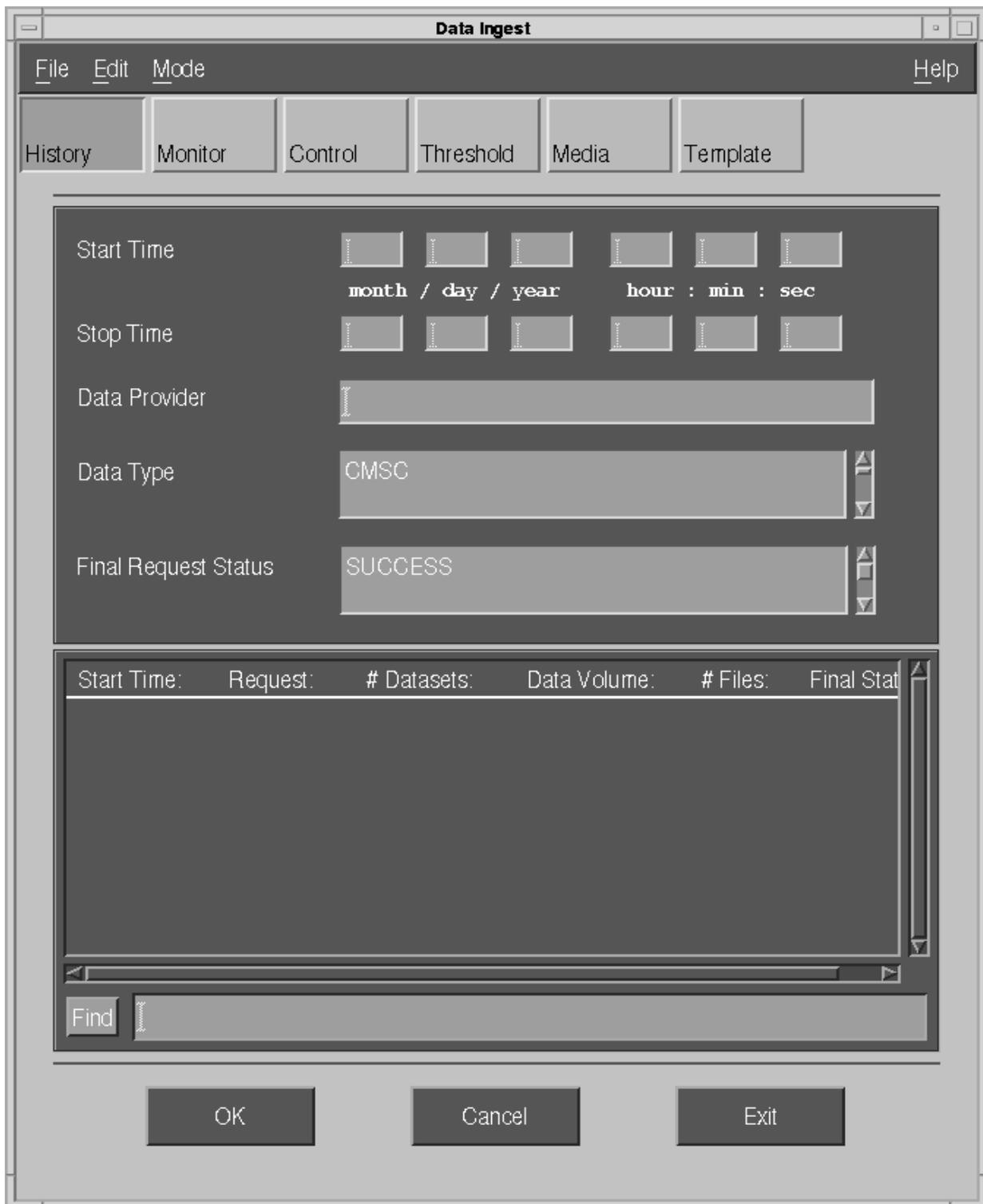
**Figure 3.9.1.6-1. Main Ingest GUI Screen**



**Figure 3.9.1.6-2. Ingest Status Monitoring GUI Screen**



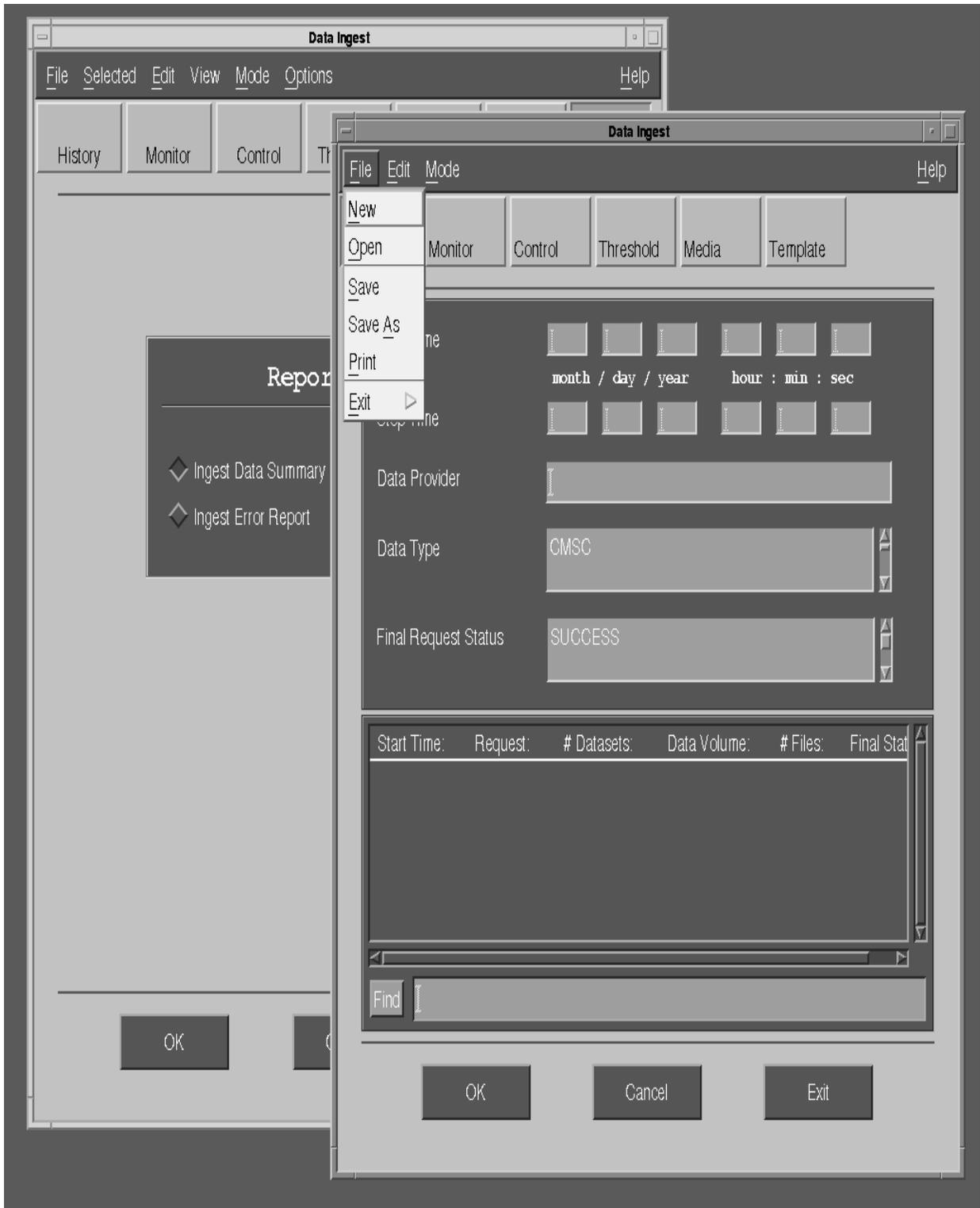
**Figure 3.9.1.6-3. Ingest Request Cancellation GUI Screen**



**Figure 3.9.1.6-4. Ingest History Log GUI Screen**



**Figure 3.9.1.6-5. Ingest Summary Report GUI Screen**



**Figure 3.9.1.6-6. Ingest Summary Report GUI Selection Screen**

## 3.9.2 TRMM Data Ingest "Fault" Scenario

### 3.9.2.1 Scenario Description

ECS Ingest will receive TRMM data from two external data sources, TSDIS and SDPF.

When the TRMM data is available, the external data source will use the ECS Automated Network Ingest to initiate the ingest processing.

When the TRMM data is available for ingest, the external data source sends a Data Availability Notice (DAN) to Ingest. Ingest validates the DAN and returns a Data Availability Acknowledgment (DAA) indicating validity of the request. If the DAN is valid, the ingest processing begins. The processing consists of: 1) transfer of the data to ECS, 2) preprocess the data, and 3) send an insert request to the Data Server for data archive. When the ingest processing completes, Ingest sends a Data Delivery Notice (DDN) informing the external data source of the completion status. The external data source responds with a Data Delivery Acknowledgment (DDA) to complete the transaction. Viewing of the messages and errors is accomplished through a combination of the information supplied in the Ingest GUI screens and that provided in the MSS-supplied event and history logs.

The Automated Network Ingest Protocol is an application-program to application-program interface. Under nominal conditions, the data ingest activities can be fully automated without the Ingest Technicians' intervention. However, there are some anomalous circumstances (e.g., network failure) where Ingest Technician involvement may be required. In this scenario, we will walk through the TRMM data ingest processing and address all the possible "fault" conditions where the Ingest Technician may be involved.

The following table list out all the fault conditions covered in this scenario.

Exceed max. number of session threshold limit	<u>Preprocessing errors:</u> – Missing required keyword
Invalid DAN contents	– Unknown data type or file type
Exceed max. number of concurrent request threshold limit	– Template out-of-sync – File type specific processing failed
Insufficient disk space	– Metadata validation error
Data expiration date/time expired	– Missing optional file
Data message sent error due to communications problem	Error getting data server name from the Advertising Service
File transfer (ftp) error	Data insert error

### **3.9.2.2 Frequency**

Every day, the ECS Ingest will be receiving approximately 6 to 12 DANs (every 2 to 4 hours) from the TSDIS interface and 1 to 5 DANs from the SDPF interface. "Fault" conditions will occur frequently during system testing, and generally less frequently thereafter. In particular, after the first six months of the TRMM mission shakeout, data-related errors will occur infrequently while resource related error may occur more frequently.

### **3.9.2.3 Assumptions**

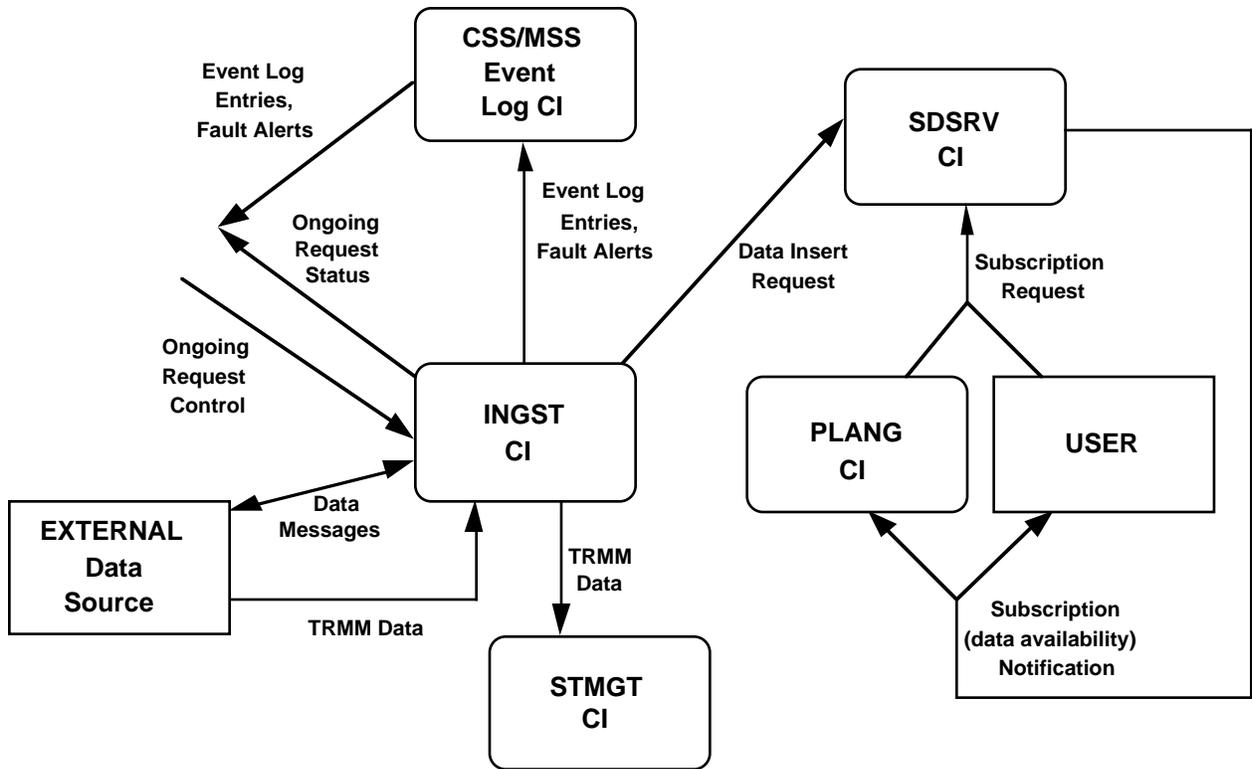
The assumptions for the scenario include:

1. The external data provider is an authorized data provider to ingest data into ECS system.
2. The external data provider has software that is capable of sending and receiving data messages (e.g., DAN, DAA) to/from the ECS Ingest software and that the data message interface convention conforms to the ICD between the ECS and the corresponding external data source.

### **3.9.2.4 Components**

The ECS Ingest software for the TRMM mission interfaces with the external data provider software to receive and send data messages, with the CSS/MSS Subsystems to log and report errors/faults and for browsing error/status logs, and with the Data Server Subsystem to insert data into the archive repository. Figure 3.9.2.4-1 depicts all the entities (both external and internal within ECS) with which ECS Ingest interacts.

Refer to "Data Insertion Scenario (nominal)" in Section 3.9.1 for details on data insertion and data subscription, if the TRMM data has been requested via subscription service.



**Figure 3.9.2.4-1. Ingest Interface Diagram**

### 3.9.2.5 Preconditions

The application-programs running under the external data provider system and the ECS Ingest system are both in a stable operations condition.

### 3.9.2.6 Detailed Steps of Process

In general, if the received DAN is not valid (e.g., missing required information), the DAA returned from ECS Ingest should be the end of the process thread. If the DAN is valid but the ingest processing was unsuccessful, the external data provider software should receive a DDN indicating the reason of ingest failure and should return a DDA to the ECS Ingest software.

Table 3.9.2.6-1 describes the processing for TRMM Data Ingest "Fault" Scenario. Based on the condition the processing is in, the steps may not be continuous. To make the table easier to read, the anomalous conditions are high-lighted in gray.

The "User" column refers to the external data provider software, the "Operator" column refers to the ECS Ingest Technician and the "ECS System" column refers to the ECS Ingest software.

**Table 3.9.2.6-1. TRMM Data Ingest "Fault" Scenario Process (1 of 8)**

<b>Step</b>	<b>Time</b>	<b>User</b>	<b>Operator</b>	<b>ECS System</b>	<b>Associated Figure</b>
1	< 5 sec	Send a DAN to ECS requesting data ingest service.		Establish an Ingest Session upon receipt of a DAN and log in the event log indicating the new ingest session connection.	
2	< 5 sec			If maximum number of Sessions* have already been established, log the error in the event log, alert the Operator, and return a DAA indicating that the system is full and an attempt should be retried again later. NOTE: based on the modeled transaction load, this condition is expected to occur very rarely, if at all.	
3	10+ min		The session connection limitation alert appears on the Operator's screen.** May increase the session threshold value using the Ingest Threshold GUI (see Figure 3.9.2 5).		
4	< 5 sec			If the number of existing sessions is within the threshold limit*, log in the event log indicating the start of DAN validation, and validate the DAN.	
5	< 5 sec			If the DAN is invalid, log the error in the event log, alert the Operator, and return a DAA indicating that the DAN is invalid.	
6			The invalid DAN alert appears on the Operator's screen.**		

Step	Time	User	Operator	ECS System	Associated Figure
7	< 5 sec			If maximum number of existing ingest requests in the system* have reached the limit, log the error to the event log, alert the Operator, and return a DAA indicating that the system is full and an attempt should be retried again later.	
8	10+ min		The request threshold limitation alert should appear on the Operator's screen.** May increase the request threshold value using the Ingest Threshold GUI (see Figure 3.9.2 2).		
9	< 5 sec	Receive DAA from ECS.		If the DAN is valid, log in the event log the receipt of valid DAN, and send a DAA indicating valid DAN is received.	
10	< 5 sec			Set a time-out timer for the disk allocation request and request disk space allocation from the Data Server.	
11	< 1 min			In the event that the Data Server has insufficient disk space, the time-out timer will expire. The Ingest will notify the Operator that the Ingest request is waiting on Data Server disk allocation.	

Step	Time	User	Operator	ECS System	Associated Figure
12	10+ min		<p>The disk allocation waiting alert appears on the Operator's screen.**</p> <p>It is the Operator's discretion whether or not to continue to wait for the disk allocation or to cancel the request.</p> <p>May monitor the status display showing subsequent ingest request processing (see Figure 3.9.2-3). and, if applicable, may cancel requests via the Ingest Control GUI (see Figure 3.9.2-4)</p>		
13	< 5 sec			If disk space is allocated, cancel the disk allocation timer, log in the event log indicating the start of data transfer, and transfer data from external location to ECS working storage.	
14	< 5 sec			If data is unavailable and the time period during which data will remain available has expired, log the error in the event log, send a DDN indicating expiration date/time exceeded and alert the Operator. Return to step 1.	
15	10+ min		<p>The expiration date/time exceeded alert appears on the Operator's screen**.</p> <p>Notify the external data provider.</p>		

16	< 5 sec			If data is not found and the time period during which data will remain available is not expired, log the error in the event log and send a DDN indicating data not found. Return to step 1.	
----	---------	--	--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

Step	Time	User	Operator	ECS System	Associated Figure
17	5+ min			<p>If ftp error, log in the event log indicating data transfer retry, and retry data transfer at a specified periodicity*.</p> <p>After N data transfer unsuccessful retries, log the error in the event log, send a DDN indicating ftp failure and alert the Operator. Return to step 1.</p>	
18	10+ min		<p>The ftp failure alert appears on the Operator screen**.</p> <p>Investigate the problem. Query the event log to determine if other communications-related failures have occurred. If needed, consult with the external data source facility to resolve the problem.</p>		
19	< 1 min			<p>If the data transfer is successful, log in the event log indicating the data transmission status and the start of the data preprocessing, and perform data preprocessing (i.e., data conversion, reformatting, and/or metadata extraction) as required for the specific data being ingested.</p>	
20	< 5 sec			<p>If in data preprocessing that 1) missing required metadata, 2) unknown data type, 3) missing required files, 4) invalid file type 5) preprocessing templates out-of-sync, or 6) the file type specific processing failure error occurred, log the error in the event log, send a DDN indicating preprocessing error, and alert the Operator. Return to step 1.</p>	

Step	Time	User	Operator	ECS System	Associated Figure
21	10+ min		<p>The appropriate data preprocessing failure alert appears on the Operator's screen**.</p> <p>Evaluate the error off-line.</p> <p>For the missing required metadata error, request the external data source to re-ingest after correction.</p> <p>For unknown data type or file type error, communicate with the external data source.</p> <p>For the out-of-sync templates or unavailable file type specific processing error, generate a trouble ticket.</p>		
22	< 5 sec			<p>If, in data preprocessing, the error indicates: 1) metadata validation error, or 2) missing optional data files, based on the instruction indication by the preprocessing templates (instituting DAAC policy), if the templates indicate to continue inserting the data, then log the error in the event log, flag the data as "bad", generate a data server insert request to store the data, and alert the Operator. Otherwise, if the templates indicate to reject the data, then log the error in the event log and send a DDN indicating preprocessing failure. Return to step 1.</p>	

Step	Time	User	Operator	ECS System	Associated Figure
23	10+ min		If the templates indicates to continue for the identified errors, the appropriate data preprocessing failure alert for each data granule appears on the Operator's screen**. Generate a Metadata Problem Report.		
24	< 5 sec			If data preprocessing is successful, log in the event log indicating retrieving data server name, and retrieve the appropriate data server name from the Advertising Service.	
25	< 5 sec			If error getting data server name, log the error in the event log, alert the Operator, and send a DDN indicating couldn't locate appropriate data server for archival.	
26			The data name retrieval failure alert appears on the Operator's screen.**		
27	< 5 sec			If data server name is retrieved successfully, log in the event log indicating inserting data, and generate a data insert request to store data and metadata into the appropriate data server.	
28	< 5 sec			If an error occurs during data insert, log the error in the event log, alert the Operator and send a DDN indicating data insert error.	
29			The data insert failure alert appears on the Operator's screen.**		

30	< 5 sec			If insert data is successful, log in the event log the successful ingest completion status, and send the DDN indicating completion status.	
----	---------	--	--	--------------------------------------------------------------------------------------------------------------------------------------------	--

Step	Time	User	Operator	ECS System	Associated Figure
31	5+ min			<p>If error sending DDN due to communication network problem, log the error in the event log indicating resending DDN due to network problem, alert the Operator, and resend the DDN.</p> <p>After N number of unsuccessful DDN retries*, log the error in the event log, and alert the Operator.</p>	
32	10+ min		<p>The communication error alert appears on the Operator's screen.****</p> <p>Contact and notify the external data source of the DDN problem .</p>		
33	< 5 sec	Receive DDN. Return a DDA to acknowledge the receipt.			
34	< 5 sec			<p>If DDA is not received within the expected time period*, log in the event log indicating resending DDN due to expected DDA not received, and resend the DDN.</p> <p>After N number DDN retries* and DDA is still not received, log the error in the event log, alert the Operator, and perform appropriate request cleanup.</p>	
35	10+ min		See step #32.		
36	< 5 sec			<p>Receive DDA from external data provider. Log in the event log indicating cleaning up the request. Perform appropriate request cleanup.</p>	

Step	Time	User	Operator	ECS System	Associated Figure
37	10+ min		<p>May bring up the Ingest history log GUI to get information on the completed request (see Figure 3.9.2-5).</p> <p>Based on the ingest Request ID stored as part of the Ingest History Log, the full event log (status/error) history of the request may be obtained by means of MSS services</p>		

\* = Ingest operator tunable parameter.

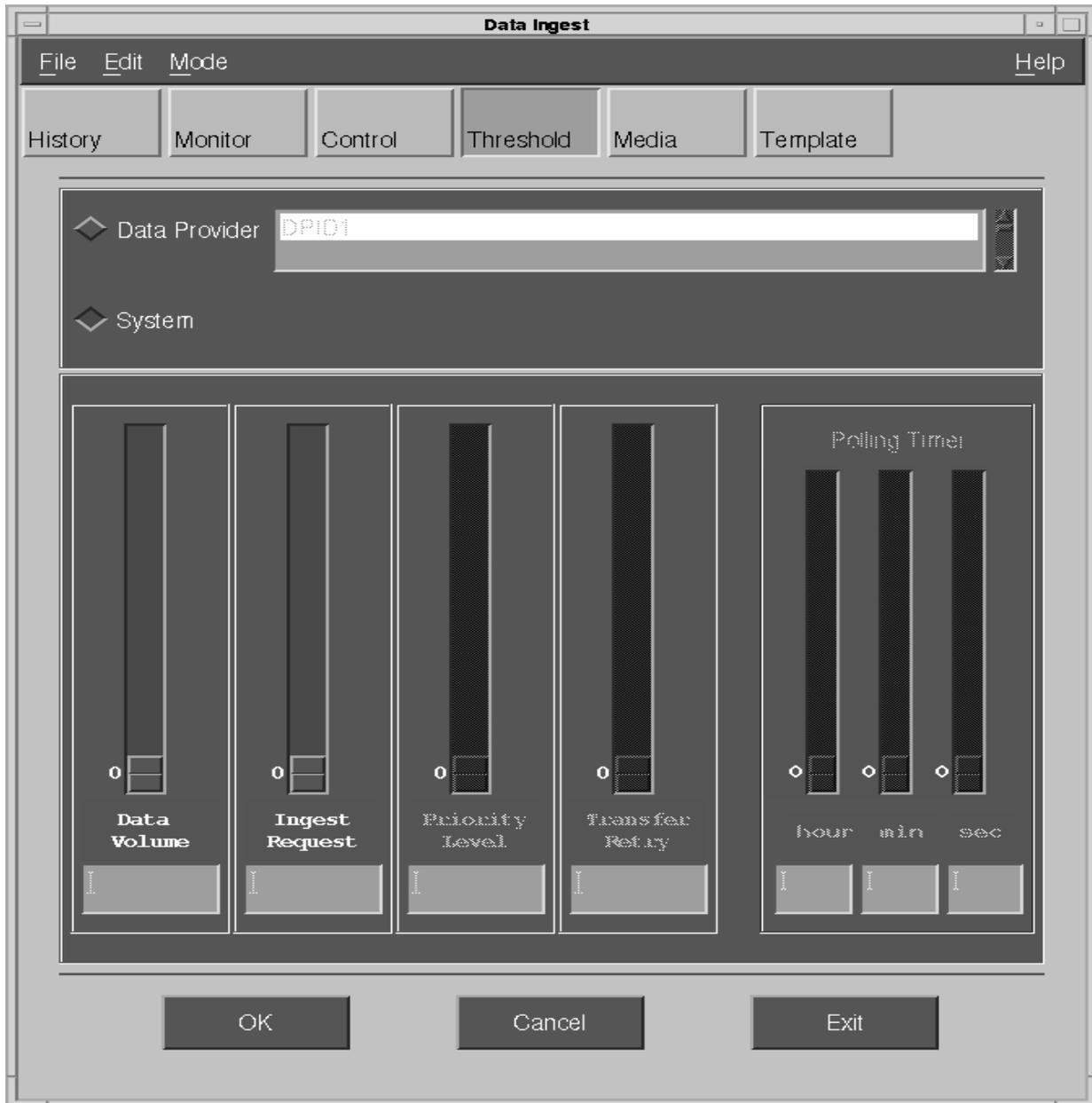
\*\* = The Operator alert service will be provided by the MSS.

\*\*\* = Storage Management operator tunable parameter

\*\*\*\* = Refer to the MSS "Production Failure Scenario" in Section 3.3.1 for details on communications failure.

### 3.9.2.7 Postconditions

If data is successfully archived into the repository, the data files on the staging should be removed and no Ingest Technician involvement is needed. However, if the ingest processing was unsuccessful, depending upon the error condition, the Ingest Technician would investigate the problem and, if needed, contact the external data source to get the problem resolved.



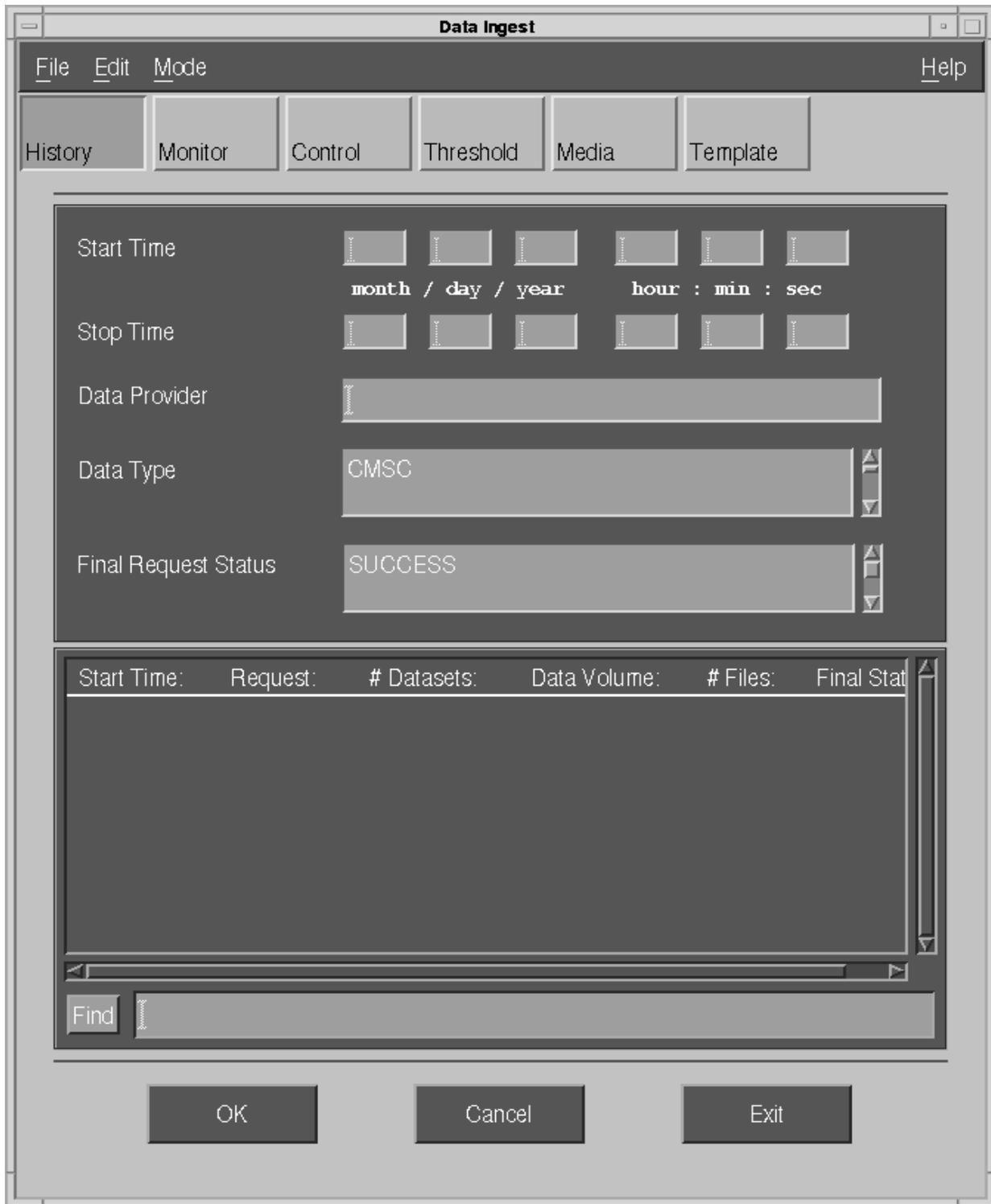
**Figure 3.9.2.6-1. Ingest Threshold Control GUI Screen**



**Figure 3.9.2.6-2. Ingest Status Monitoring GUI Screen**



**Figure 3.9.2.6-3. Ingest Request Control GUI Screen**



**Figure 3.9.2.6-4. Ingest History Log GUI Screen**

### **3.9.3 TRMM Ancillary Data Ingest Scenario**

The following scenario discusses the operations staff involvement in the ingest of ancillary data required for processing of TRMM data. This scenario will be implemented at the GSFC, LaRC, and MSFC DAACs.

The Data Ingest Technician is the primary operations staff member involved in the automated ancillary data ingest process.

#### **3.9.3.1 Scenario Description**

The following scenario is assumed to occur during a given day of the TRMM Mission (Release A) period at the GSFC, LaRC, and MSFC DAACs. Three subscenarios are to be discussed:

- Data Ingest Technician monitoring of ongoing ingest request processing (including the ancillary ingest request processing)
- Data Ingest Technician cancellation of an ingest request
- Data Ingest Technician viewing of summary ingest request processing status and statistics

The GSFC, LaRC, and MSFC DAACs expect to receive different TRMM ancillary products on a daily, weekly, and monthly basis. The size of the delivery at the sites varies from a few megabytes per month to over 100 megabytes per day, depending on the data product. Delivery is accomplished electronically through the EOSDIS Backbone Network (EBnet).

Note: in the event of a long-term failure of the network interface between the ancillary data providers and ECS, an 8mm tape backup mechanism may be available for certain interfaces. The hard media ingest scenario (3.9.4) discusses operations staff procedures for that contingency scenario.

#### **3.9.3.2 Frequency**

In the nominal case the deliveries are expected at regular times during the staffed portion of the day. However, the system is set up to accept deliveries at any time of the day, including unstaffed periods. Different ancillary products are anticipated to be delivered with frequencies of 4 times per day, once per day, once per week, and once per month.

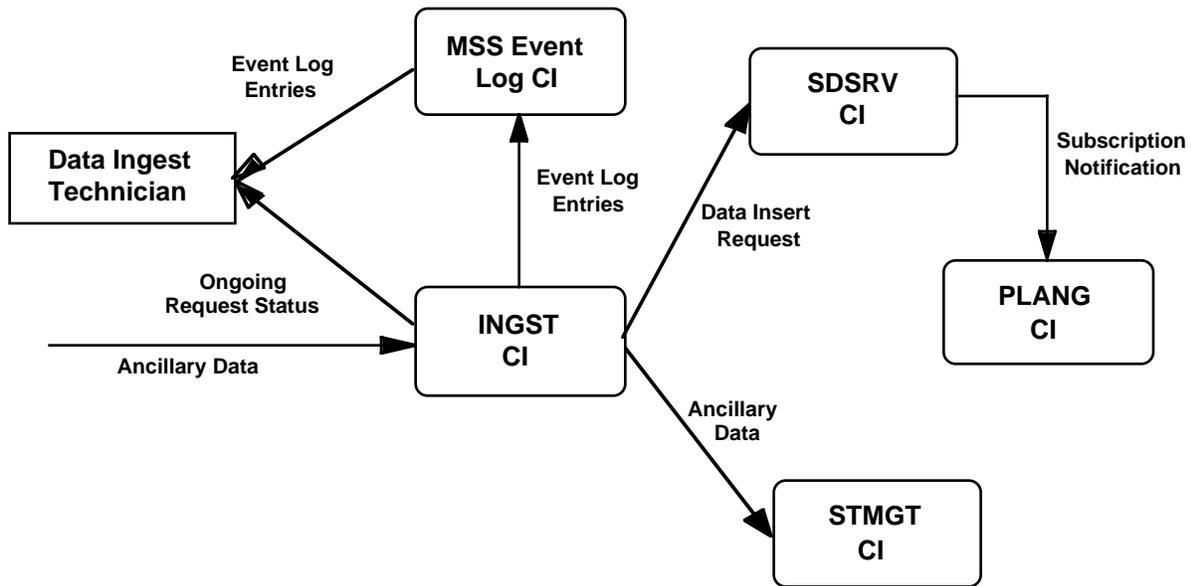
#### **3.9.3.3 Assumptions**

Assumptions underlying this scenario are as follows.

1. The system at the DAAC is in stable operations.
2. The nominal TRMM ancillary data ingest process is discussed in this scenario. "Fault" conditions are discussed in the TRMM data ingest fault scenario (3.9.2).

### 3.9.3.4 Components

There are five configuration items (CIs) involved with this scenario. Figure 3.9.3.4-1 indicates the interaction between the operations personnel, the INGST CI, the TBD MSS Event Log CI, the SDSRV CI, the STMGT CI, and the PLANG CI.



**Figure 3.9.3.4-1. TRMM Ancillary Ingest CI/Operations Staff Interaction**

### 3.9.3.5 Preconditions

The following preconditions are assumed for this scenario:

1. Data and metadata configuration information has been implemented in ECS tables to appropriately describe the structure of incoming ancillary data and the processing to be performed on the data.
2. All necessary technical and management personnel have been trained in the use of MSS event log and INGST status monitoring tools.

### 3.9.3.6 Detailed Steps of Process

Table 3.9.3.6-1 represents the details of the ancillary data ingest activities. The time scales indicated are approximate. The "User" in this case represents the external data provider. The "Operator" is the Data Ingest Technician, who is responsible for monitoring ongoing data ingest processing.

**Table 3.9.3.6-1. RMM Ancillary Data Ingest Scenario Process (1 of 3)**

<b>Step</b>	<b>Time</b>	<b>User</b>	<b>Operator</b>	<b>System</b>	<b>Associated Figure</b>
0	Prior to start of TRMM mission		The TBD operations staff enters a subscription requesting notification upon receipt of TRMM ancillary data.	The System (SDSRV CI) stores the subscription pending receipt of TRMM ancillary data.	See TBD scenario for subscription setup GUI.
1	Start of work day		The Data Ingest Technician invokes the Ingest Status Monitor tool (GUI display) from the main Ingest GUI screen.	The System identifies ongoing ingest requests (stored in Sybase tables) and displays them.	3.9.3.6-1
2	Ancillary data available	Ancillary data provider application software automatically sets up to write ancillary data and Delivery Record to specified location.		The System automatically checks a predetermined network location for the presence of a Delivery Record file. Once a Delivery Record file is located, the System automatically coordinates the ancillary data transfer with the data provider using ftp services. After the ancillary data and Delivery Record file is received, the INGEST CSCI automatically checkpoints request information extracted from the Delivery Record into a Sybase data base.	

3	Ancillary data available + < a few minutes		The Data Ingest Technician periodically reviews Ingest Status Monitor display. The Technician looks for ingest requests that have been queued for an unexpected period (in the TRMM Mission era, anything on the queue for more than a few minutes.	The System automatically extracts metadata from transferred ancillary data and checks the metadata (e.g., range checks). Format conversion (e.g., GRIB to HDF-EOS for NMC-ETA data) is automatically performed for all ancillary products previously identified as requiring conversion. It then inserts the data and metadata into the appropriate Data Server. Request state (active, file transferred, data insertion complete, etc.) is updated in the checkpointed request information.	3.9.3.6-2
---	--------------------------------------------	--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------

<b>Step</b>	<b>Time</b>	<b>User</b>	<b>Operator</b>	<b>System</b>	<b>Associated Figure</b>
4	Ancillary data available + < a few minutes		The Data Ingest Technician may cancel a request if desired.	The System accepts cancellation request based on a given request ID.	3.9.3.6-3
5	Ancillary data available + < a few minutes		The Data Ingest Technician periodically reviews the MSS Event Log to visually determine anomalous conditions (e.g., a pattern of metadata check errors).	The System automatically logs events by means of the MSS Event Logging capability. "Events" include detection of out-of-range metadata values, incompletely-transferred data files, etc. Based on DAAC policy, selected events may be identified as "alerts", which trigger a visual change of state at the MSS Event Log display.	See scenario TBD for a discussion of MSS Event Log displays
6	Ancillary data available + a few minutes		The Data Ingest Technician observes the removal of the completed ingest request from the Status Monitor display.	Upon completion (successful or unsuccessful) of data insertion into the Data Server, status is automatically returned to the data provider by means of electronic mail. Summary information is retained in the Sybase data base (as Ingest History Log data).	
7	Ancillary data available + a few minutes			Upon completion (successful or unsuccessful) of data insertion into the Data Server, the SDSRV CI automatically determines the existence of subscriptions on the receipt of the TRMM ancillary data. A subscription notice is sent to the requesting entity.	

Step	Time	User	Operator	System	Associated Figure
8	After ingest request completion.		<p>The Data Ingest Technician may view summary information about completed ingest requests using the GUI Ingest History Log tool.</p> <p>The Data Ingest Technician may generate a summary report on completed ingest requests. The report (in two parts) gives summary statistics (e.g., number of data granules ingested, data volume ingested) and error statistics (e.g., number of errors of a given type encountered) for a specified time range.</p>	<p>The System provides access to Sybase data base tables containing summary information on completed ingest requests, including completion status, data volume ingested, etc.</p>	3.9.3.6-4
9	After ingest request completion.		<p>The Data Ingest Technician may generate a summary report on completed ingest requests. The report (in two parts) gives summary statistics (e.g., number of data granules ingested, data volume ingested) and error statistics (e.g., number of errors of a given type encountered) for a specified time range.</p>	<p>The System provides access to Sybase data base tables containing summary information on completed ingest requests, including completion status, data volume ingested, etc.</p>	3.9.3.6-5 and 3.9.3.6-6

### 3.9.3.7 Postconditions

At the completion of the above scenario, the Planning Subsystem (PLANG) is notified by SDSRV of the availability of the data. The Processing Subsystem (PRONG) is then invoked by the PLANG once all Level 0 data and associated ancillary data are available. PRONG will retrieve data from the Level 0 Data Server and process to higher levels utilizing the ingested ancillary data per scenario 3.13.1.



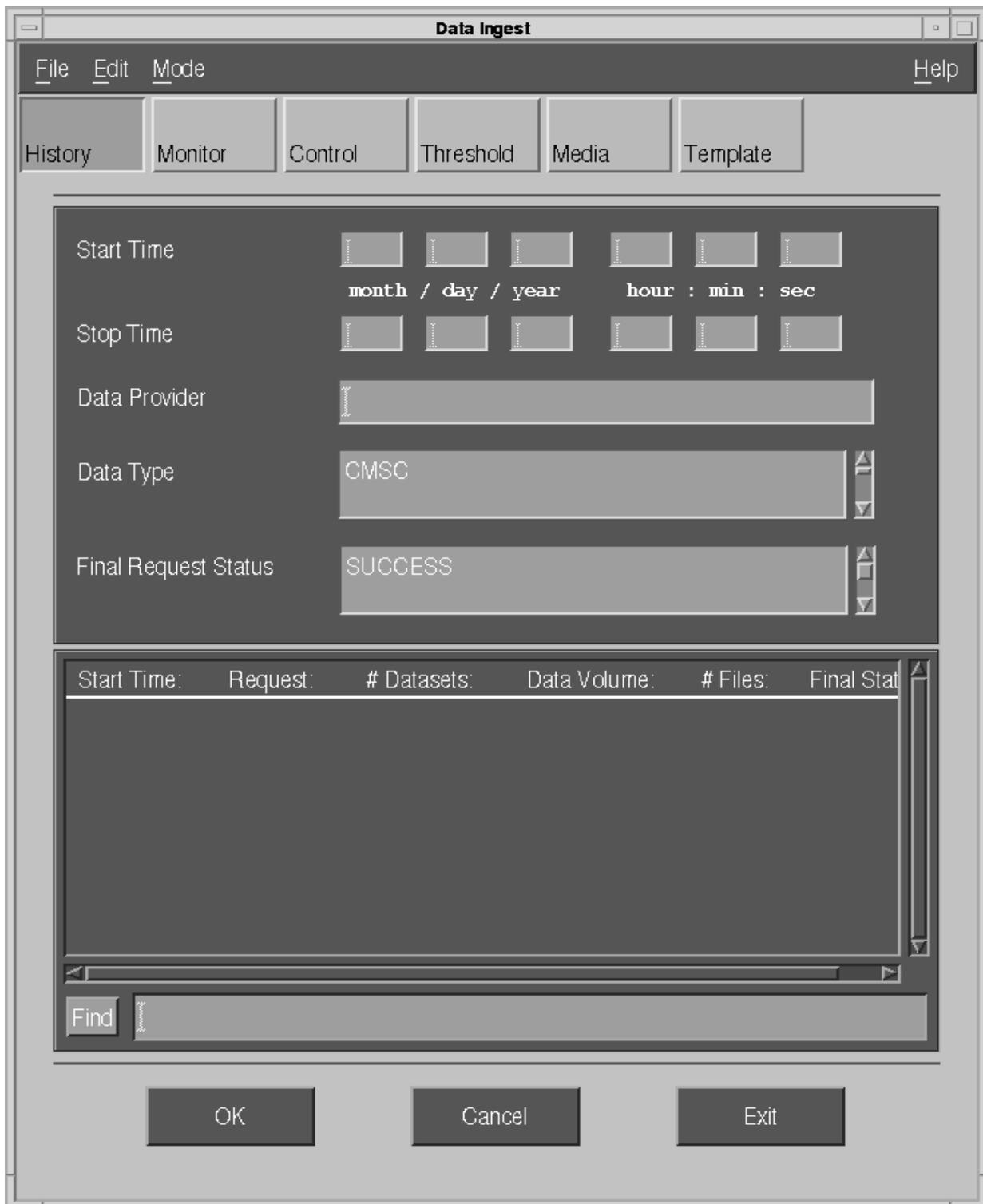
**Figure 3.9.3.6-1. Main Ingest GUI Screen**



**Figure 3.9.3.6-2. Ingest Status Monitoring GUI Screen**



**Figure 3.9.3.6-3. Ingest Request Cancellation GUI Screen**



**Figure 3.9.3.6-4. Ingest History Log GUI Screen**



**Figure 3.9.3.6-5. Ingest Summary Report GUI Screen**



**Figure 3.9.3.6-6. Ingest Summary Report GUI Selections Screen**

## **3.9.4 Hard Media Ingest Scenario**

### **3.9.4.1 Scenario Description**

This scenario describes the processing required for the ingest of hard media. The media that is received at the DAAC is checked for readiness to ingest. The ingest technician compares the received media to a media ingest readiness checklist and invokes the ingest client s/w via the GUI once the media has been readied. A device allocation is requested from the Data Server peripheral pool. The ingest technician receives the device id and is prompted to mount the media. The existence of a Delivery Record file describing the media contents is checked and a summary of the Delivery Record contents is logged. The data is transferred from media to working storage, and basic metadata extraction and validation is performed. The ingest function then generates a data server insert request to store the data and metadata in the appropriate data repository. E-mail notification of successful ingest is sent to the data provider if a network address is available.

### **3.9.4.2 Frequency**

Other than the SCF and ASTER interface, there is no routine schedule for the media ingest processing. The media ingest service is generally performed on an ad hoc or as required basis.

- SCF is expected to use the media ingest service twice a month for ingesting the Data Production Software Delivery Package into the ECS system.
- ASTER is expected to send media to ECS on a weekly basis.
- For the SDPF, the media ingest service will be used when there is a network communication failure. The service is expected to be used very infrequently by SDPF since 1) network communications failures do not occur often and 2) when a network communication failure happens, the duration of the network failure time will be shorter than the time period for which the data will be kept in the storage system. The SDPF will retain their data on the magnetic disk for up to 5 days.
- The frequency of media ingest for the V0 data migration will be investigated as part of the V0 migration facility analysis task. V0 migration may provide the bulk of the hard media ingest load at some of the DAACs. This will be quantified as part of the V0 data migration analysis currently underway.

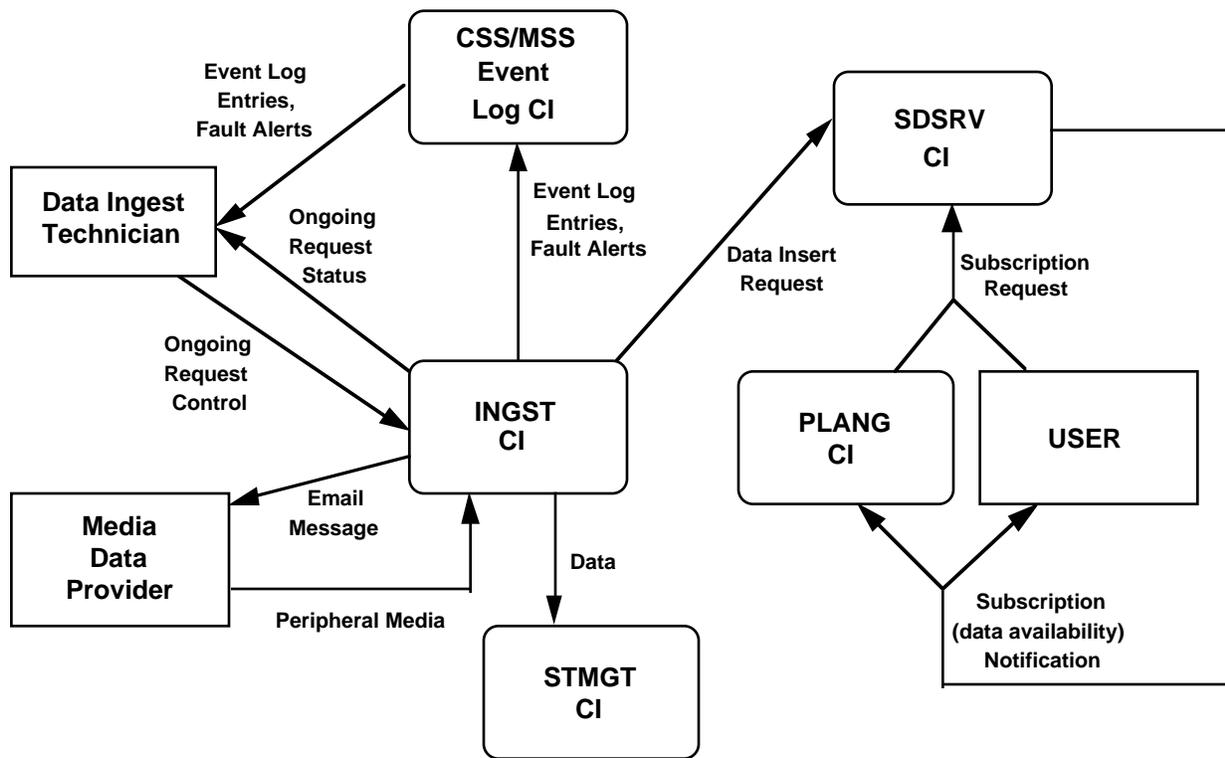
### **3.9.4.3 Assumptions**

It is assumed that the media provider is an authorized user for data ingest. That is, the media provider has an ingest account registered through the DAAC System Administrator.

### **3.9.4.4 Components**

The media ingest process interfaces with the CSS/MSS Subsystems to log and report errors/faults and for browsing error/status logs, and with the Data Server Subsystem to insert data into the archive repository. Figure 3.9.4.4-1 depicts all the entities with which ECS Ingest interacts.

Refer to "Data Insertion Scenario (nominal)" in Section 3.9.1 for details on data insertion and data subscription, for data that has been requested via subscription service.



**Figure 3.9.4.4-1. Media Ingest Diagram Scenario Components**

### 3.9.4.5 Preconditions

Prior to initiating the media ingest, the Ingest Technician needs to perform the appropriate readiness checks. This includes ensuring that a packaging slip is enclosed and that the number of media in the package matches the number indicated on the packaging slip. The packaging slip should contain information (e.g., number of media in the package, data type associated with the data on the media) the Ingest Technician needs to fill in all the necessary fields of the GUI interface for media ingest.

### 3.9.4.6 Detailed Steps of Process

Table 3.9.4.6-1 describes the processing for media ingest. Based on the condition the processing is in, the steps may not be continuous. To make the table easier to read, the anomalous conditions are high-lighted in gray.

In the table, the "Operator" column refers to the Ingest Technician and the "ECS System" column refers to the ECS Ingest System.

**Table 3.9.4.6-1. Media Ingest Scenario Process (1 of 5)**

Step	Time	User	Operator	ECS System	Associated Figure
1	< 30 min	User packages the hard media and delivers the package to the DAAC.	(The hard media is received at the DAAC.) Compare against the media ingest readiness checklist, verify user authenticity based on information obtained from the data provider's user profile, and invoke the ingest client s/w via the GUI (see Figure 3.9.4-2).		
2	< 5 sec			Set a time-out timer for the media device allocation request and request media device allocation from the Data Server peripheral pool.	
3	< 1 min			In the event that the Data Server has no device available, the time-out timer will expire. The Ingest will notify the Operator that the Ingest request is waiting on Data Server device allocation.	
4	10+ min		The device allocation waiting alert appears on the Operator's screen.* It is at the Operator's discretion whether or not to terminate the media device allocation request. Operator may monitor the status display showing subsequent ingest request processing (see Figure 3.9.4-4) and if applicable, cancel requests (see Figure 3.9.4-5).		

5	< 5 sec			If a device is allocated, request the mount service from the Data Server. Data Server will prompt the Ingest Technician to mount the media.	
---	---------	--	--	---------------------------------------------------------------------------------------------------------------------------------------------	--

Step	Time	User	Operator	ECS System	Associated Figure
6	< 10 min		Receive the tape mount prompt.** Mount the media according to the supplied device ID.		
7	< 5 sec			If mount error is detected, log the error on the event log indicating mount failure, notify the Operator and wait for the response.	
8	< 10 sec		Receive the mount failure prompt on the Operator's screen.* Investigate the device. If the media is not mounted properly, re-mount the tape and respond the prompt. If the device is bad, put the bad device to off-line and respond the prompt.	Upon receipt of the response: for improper mount error, if the number of retry is within the threshold limit, go to step #5 to retry the mount service. for bad device error, if the number of retry is within the threshold limit, go to step #2 to retry the allocate service. If the number of retry exceeds the threshold limit, log the error in the event log, alert the Operator and terminate the program.	
9	< 5 sec		If the number of retry exceeds the threshold limit, the media ingest failure alert appears on the Operator's screen.*		
10	< 5 sec			If the mount is successful, login the event log indicating start of data transfer, and transfer data from media to working storage.	
11	< 5 sec			If the media is bad (e.g., error reading the media), log the error in the event log, send an Email to the User indicating bad media error, and alert the Operator. The operator may choose to attempt to read the media on another device to ensure that the media is bad before sending a message to the media provider. Go to step #24.	

Step	Time	User	Operator	ECS System	Associated Figure
12	< 5 sec		The bad media alert appears on the Operator's screen.*		
13	< 5 sec			If device error is detected, log the error in the event log, notify the Operator that the device is bad and wait for the Operator's response.	
14	< 5 min		The bad device message prompt appears on the Operator's screen.* Put the bad device to off-line and respond the prompt.	Upon receipt of the response: if the number of allocate retry is within the threshold limit, go to step #2 to retry the allocate service. if the number of allocate retry exceeds the threshold limit, log the error in the event log, notify the Operator and terminate the program.	
15	< 5 sec		The media ingest failure alert appears on the Operator's screen.*		
16	< 5 sec			If the data transfer is successful, log the data transmission status in the event log.	
17	< 5 sec			Request the dismount service from the Data Server. The Data Server will prompt the Ingest Technician to dismount the media from the device.	
18	< 10 min		Receive the dismount message.** Dismount the media from the appropriate device.		
19	< 5 sec			Deallocate peripherals.	
20	< 5 sec			Check for the existence of a Delivery Record file.	

21	< 5 sec			If Delivery Record file does not exist, log the error in the event log and send an Email to the User indicating missing Delivery Record file. Go to step #24.	
----	---------	--	--	------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

Step	Time	User	Operator	ECS System	Associated Figure
22	< 1 min			<p>If Delivery Record file exist, perform data preprocessing and data insertion.</p> <p>For details on nominal processing, refer to Section 3.9.1 (TRMM L0 Data Ingest), steps 3 to 9.</p> <p>For details on anomalous condition handling, refer to steps Section 3.9.2 (TRMM Data Ingest "Fault" Scenario), steps 19 to 27.</p>	
23	< 5 sec			<p>If the Email address is available, send an Email to the User indicating the completion status for the media ingest.</p>	
24	< 10 min		<p>May bring up the Ingest history log GUI to get information on the completed request (see Figure 3.9.4-6).</p> <p>Based on the ingest Request ID stored as part of the Ingest History Log, the full event log (status/error) history of the request may be obtained by means of MSS services</p>		
25	< 5 min	The user receives the Email notification indicating the completion status of the media ingest.			
26	< 5 sec			<p>If the Email address is NOT provided, notify the Operator of the completion of media ingest.</p>	

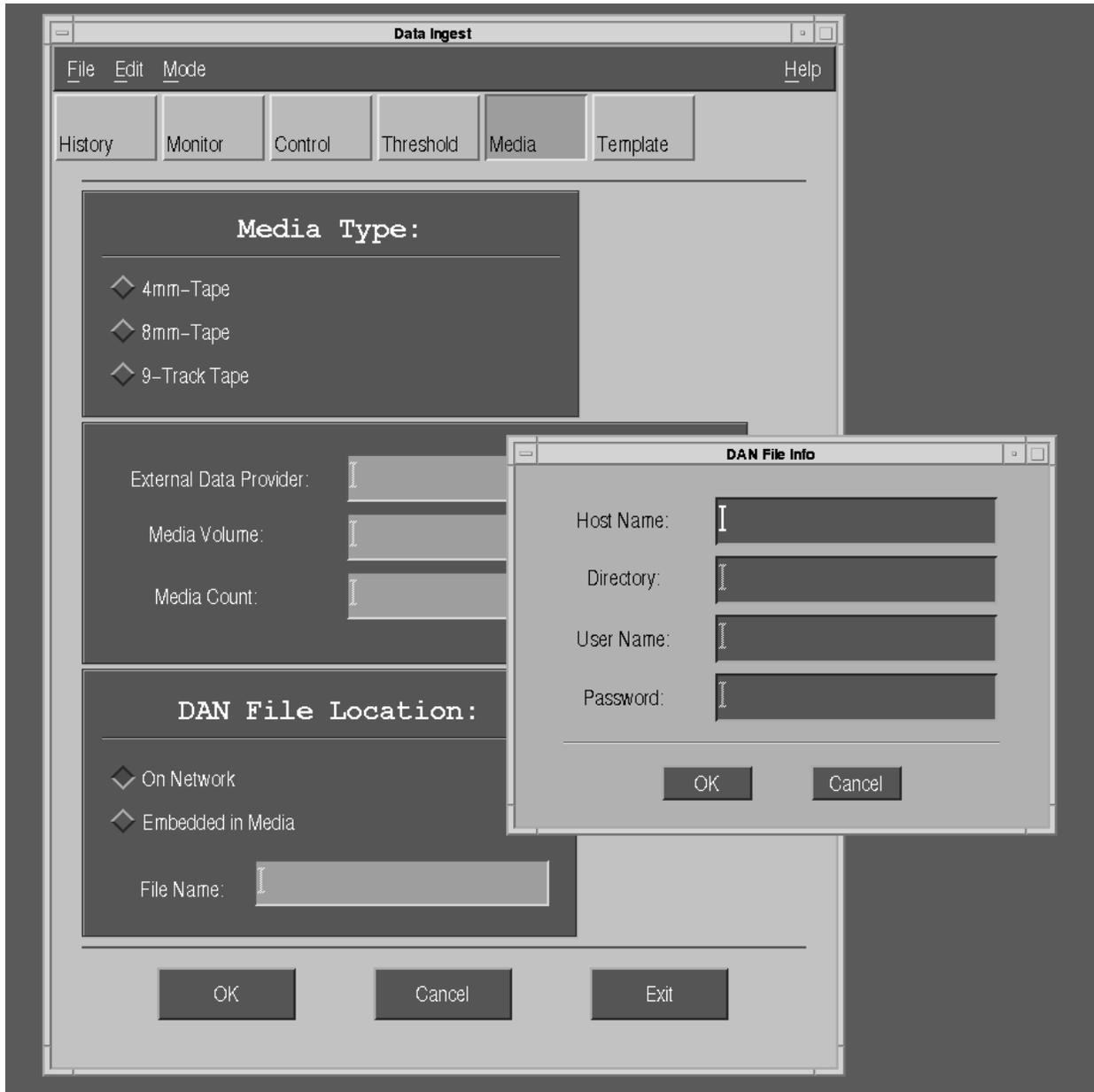
Step	Time	User	Operator	ECS System	Associated Figure
27	< 10 min		Inform the User of the media ingest completion via telephone, facsimile, or other means based on DAAC policy.		

\* = The Operator alert service will be provided by the MSS.

\*\* = The Mount and Dismount prompt service will be provided by the Data Server.

### 3.9.4.7 Postconditions

When the media ingest completes, the data in the media should be archived into the archive repository system by the Data Server Subsystem. If the data is being subscribed, the Data Server Subsystem will notify the subscriber indicating the availability of data. Lastly, the media provider should receive an E-mail notification indicating the completion status of the media ingest, if an E mail address is provided. Otherwise, the Ingest Technician will inform the media provider of the completion status via telephone.



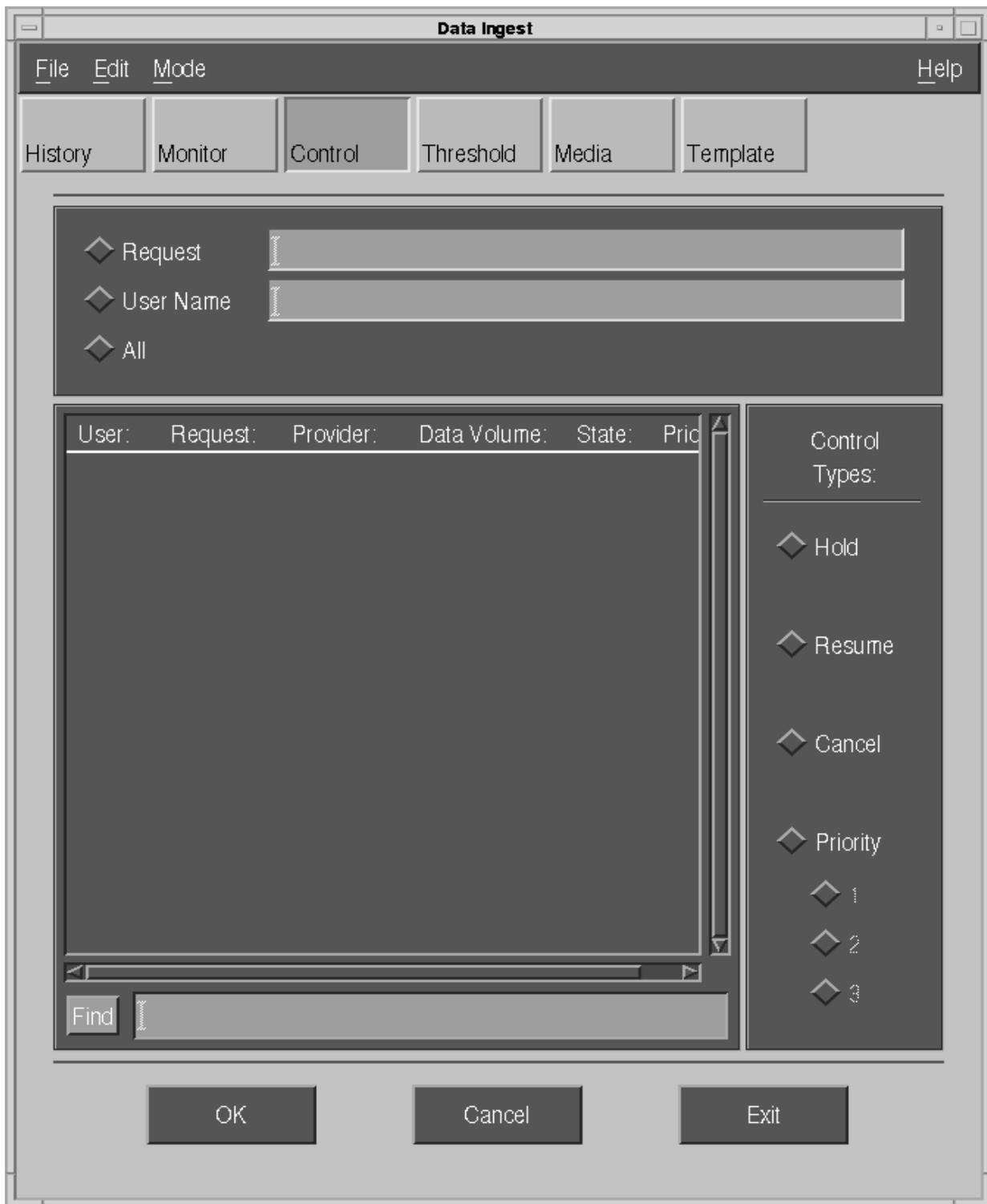
**Figure 3.9.4.6-1. Ingest Media GUI Screen**



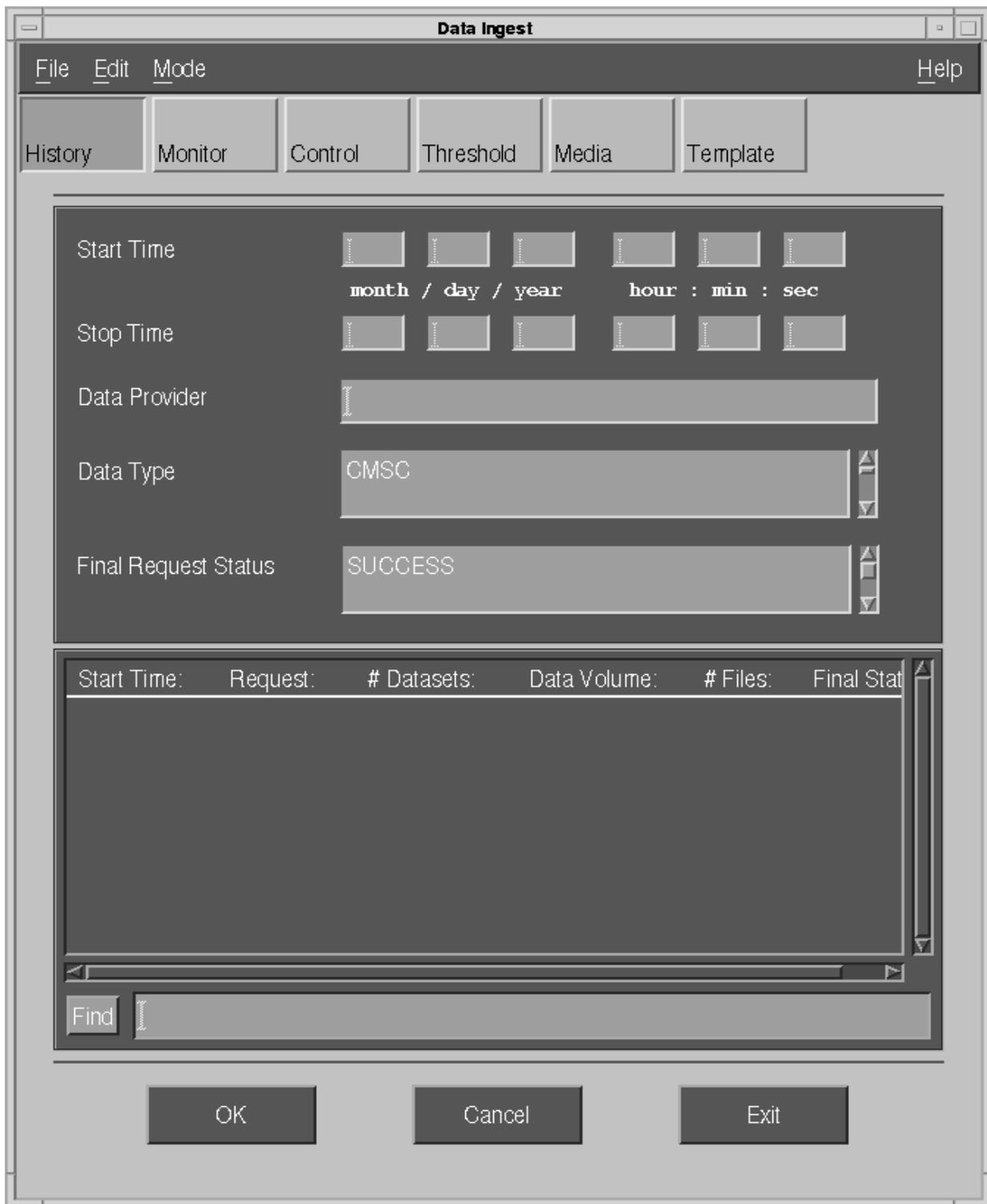
**Figure 3.9.4.6-2. Ingest Operator Action Prompt GUI Screen**



**Figure 3.9.4.6-3. Ingest Status Monitoring GUI Screen**



**Figure 3.9.4.6-4. Ingest Request Control GUI Screen**



**Figure 3.9.4.6-5. Ingest History Log**

### **3.9.5 Version 0 Data Ingest Scenario**

The following scenario discusses the operations staff involvement in the ingest of data from the Version 0 (V0) system into ECS. This scenario will be implemented at the ASF, EDC, GSFC, JPL, LaRC, MSFC, and NSIDC DAACs.

The Data Ingest Technician is the primary operations staff member involved in the Version 0 data ingest process. ECS M&O engineering staff are also involved in V0 migration activities that precede the actual ingest of the V0 data into ECS. These activities include the benchmarking of the migration software and overseeing the preprocessing of the V0 data prior to ingest. This scenario addresses the ingest of the data and population of the appropriate data servers once the V0 migration facility tasks have been completed.

#### **3.9.5.1 Scenario Description**

The following scenario is assumed to occur during a given day of the AM-1/L7 Mission (Release B) period at the ASF, EDC, GSFC, JPL, LaRC, MSFC, and NSIDC DAACs. Three subscenarios are to be discussed:

- Data Ingest Technician monitoring of ongoing ingest request processing (including the Version 0 ingest request processing)
- Data Ingest Technician cancellation of an ingest request
- Data Ingest Technician viewing of summary ingest request processing status and statistics

Release A daily migration volumes will range from approximately 630 MB to 8.3 GB per day.

In the Release B timeframe, the ASF, EDC, GSFC, JPL, LaRC, MSFC, and NSIDC DAACs expect to migrate an average of between 200 MB and 11.4 GB per day depending on the total volume to be migrated at each DAAC. Ingestion of the migrated data is accomplished either through electronic transfer over the internal DAAC network, or via hard media.

#### **3.9.5.2 Frequency**

In the nominal case migration is expected to be performed 24 hours/day, 7 days/week. Data products will be migrated as they are released from the V0 migration facility, at a frequency that is determined by the size of the data sets and complexity of conversions being performed. Migration during unstaffed periods at certain DAACs will be accommodated through the use of the Automated Network Ingest capabilities discussed in Section 3.9.1 and/or the use of tape stackers for hard media ingest.

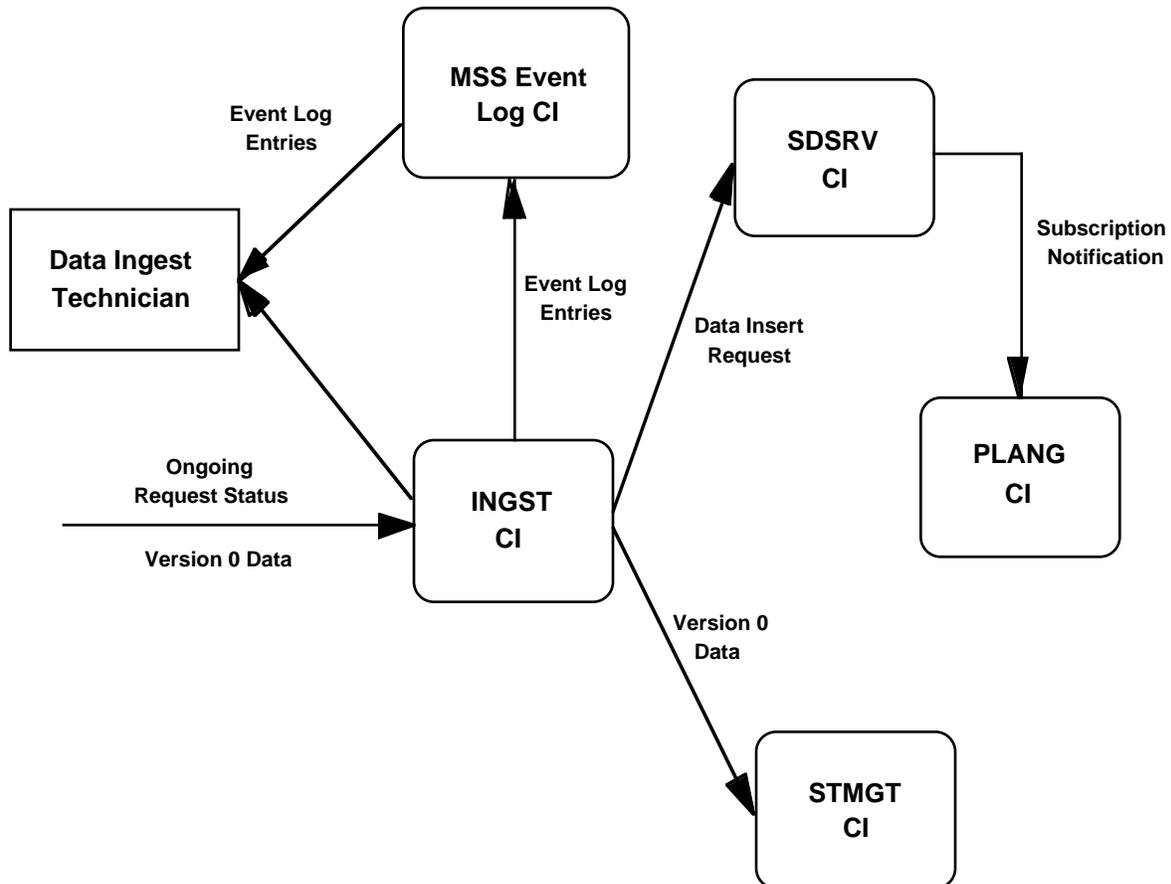
#### **3.9.5.3 Assumptions**

Assumptions underlying this scenario are as follows.

1. The system at the DAAC is in stable operations.
2. The nominal Version 0 data ingest process is discussed in this scenario. "Fault" conditions are discussed in the data ingest fault scenario (3.9.2).

### 3.9.5.4 Components

There are five configuration items (CIs) involved with this scenario. Figure 3.9.5.4-1 indicates the interaction between the operations personnel, the INGST CI, the TBD MSS Event Log CI, the SDSRV CI, the STMGT CI, and the PLANG CI.



**Figure 3.9.5.4-1. Version 0 Ingest CI/Operations Staff Interaction**

### 3.9.5.5 Preconditions

The following preconditions are assumed for this scenario:

1. Data and metadata configuration information obtained from the V0 migration facility has been implemented in ECS tables to appropriately describe the structure of incoming Version 0 data and the processing to be performed on the data.
2. All necessary technical and management personnel have been trained in the use of MSS event log and INGST status monitoring tools.

### **3.9.5.6 Detailed Steps of Process**

Table 3.9.5.6-1 represents the details of the Version 0 data ingest activities. The time scales indicated are approximate. The "User" in this case represents the Version 0 migration facility. The "Operator" is the Data Ingest Technician, who is responsible for monitoring ongoing data ingest processing. This scenario depicts the ingest of Version 0 data using the automated network ingest mechanism described in Section 3.9.1. In the event that hard media transfer of the V0 data is to be used, see the scenario described in Section 3.9.4.

**Table 3.9.5.6-1. Version 0 Data Ingest Scenario Process (1 of 3)**

<b>Step</b>	<b>Time</b>	<b>User</b>	<b>Operator</b>	<b>System</b>	<b>Associated Figure</b>
0	Prior to start of V0 migration		The TBD operations staff enters a subscription requesting notification upon receipt of specific V0 data.	The System (SDSRV CI) stores the subscription pending receipt of V0 data.	Subscription setup GUI TBD
1	Start of work day		The Data Ingest Technician invokes the Ingest Status Monitor tool (GUI display) from the main Ingest GUI screen.	The System identifies ongoing ingest requests (stored in Sybase tables) and displays them.	3.9.5.6-1
2	V0 data available	V0 migration facility application software automatically sets up to transfer V0 data. A Data Availability Notice (DAN) is sent to ECS. A Data Availability Acknowledgement (DAA) is received.		The System automatically checks the received DAN and returns a DAA. Once a request to ingest V0 data is received, the INGST CSCI automatically checkpoints request information extracted from the DAN into a Sybase data base. The System automatically coordinates the V0 data transfer with the V0 migration facility.	
3	V0 data available + < a few minutes		The Data Ingest Technician periodically reviews Ingest Status Monitor display. The Technician looks for ingest requests that have been queued for an unexpected period.	The System automatically extracts metadata from transferred V0 data, checks the metadata (e.g., range checks), and inserts the data and metadata into the appropriate Data Server. Request state (active, file transferred, data insertion complete, etc.) is updated in the checkpointed request information.	3.9.5.6-2
4	V0 data available + < a few minutes		The Data Ingest Technician may cancel a request if desired.	The System accepts cancellation request based on a given request ID.	3.9.5.6-3

Step	Time	User	Operator	System	Associated Figure
5	V0 data available + < a few minutes		The Data Ingest Technician periodically reviews the MSS Event Log to visually determine anomalous conditions (e.g., a pattern of metadata check errors).	The System automatically logs events by means of the MSS Event Logging capability. "Events" include detection of out-of-range metadata values, incompletely-transferred data files, etc. Based on DAAC policy, selected events may be identified as "alerts", which trigger a visual change of state at the MSS Event Log display.	MSS Event Log displays TBD
6	V0 data available + a few minutes	V0 migration facility application software receives a Data Delivery Notice (DDN), indicating the status of the V0 data ingest. The V0 migration facility application software transmits a Data Delivery Acknowledgment (DDA) to acknowledge receipt of the DDN.	The Data Ingest Technician observes the removal of the completed ingest request from the Status Monitor display.	Upon completion (successful or unsuccessful) of data insertion into the Data Server, status is automatically returned to the data provider (V0 migration facility) by means of the DDN. After receipt of a DDA, ongoing ingest request information is deleted. Summary information is retained in the Sybase data base (as Ingest History Log data).	
7	V0 data available + a few minutes			Upon completion (successful or unsuccessful) of data insertion into the Data Server, the SDSRV CI automatically determines the existence of subscriptions on the receipt of the Version 0 data. A subscription notice is sent to the requesting entity.	

<b>Step</b>	<b>Time</b>	<b>User</b>	<b>Operator</b>	<b>System</b>	<b>Associated Figure</b>
8	After ingest request completion.		The Data Ingest Technician may view summary information about completed ingest requests using the GUI Ingest History Log tool.	The System provides access to Sybase data base tables containing summary information on completed ingest requests, including completion status, data volume ingested, etc.	3.9.5.6-4
9	After ingest request completion.		The Data Ingest Technician may generate a summary report on completed ingest requests. The report (in two parts) gives summary statistics (e.g., number of data granules ingested, data volume ingested) and error statistics (e.g., number of errors of a given type encountered) for a specified time range.	The System provides access to Sybase data base tables containing summary information on completed ingest requests, including completion status, data volume ingested, etc.	3.9.5.6-5 and 3.9.5.6-6

### 3.9.5.7 Postconditions

At the completion of the above scenario, the Planning Subsystem (PLANG) is notified by SDSRV of the availability of the data. The Processing Subsystem (PRONG) is then invoked by the PLANG once all Version 0 data and associated ancillary data are available. PRONG will retrieve data from the Data Server and process to higher levels per scenario 3.13.1.



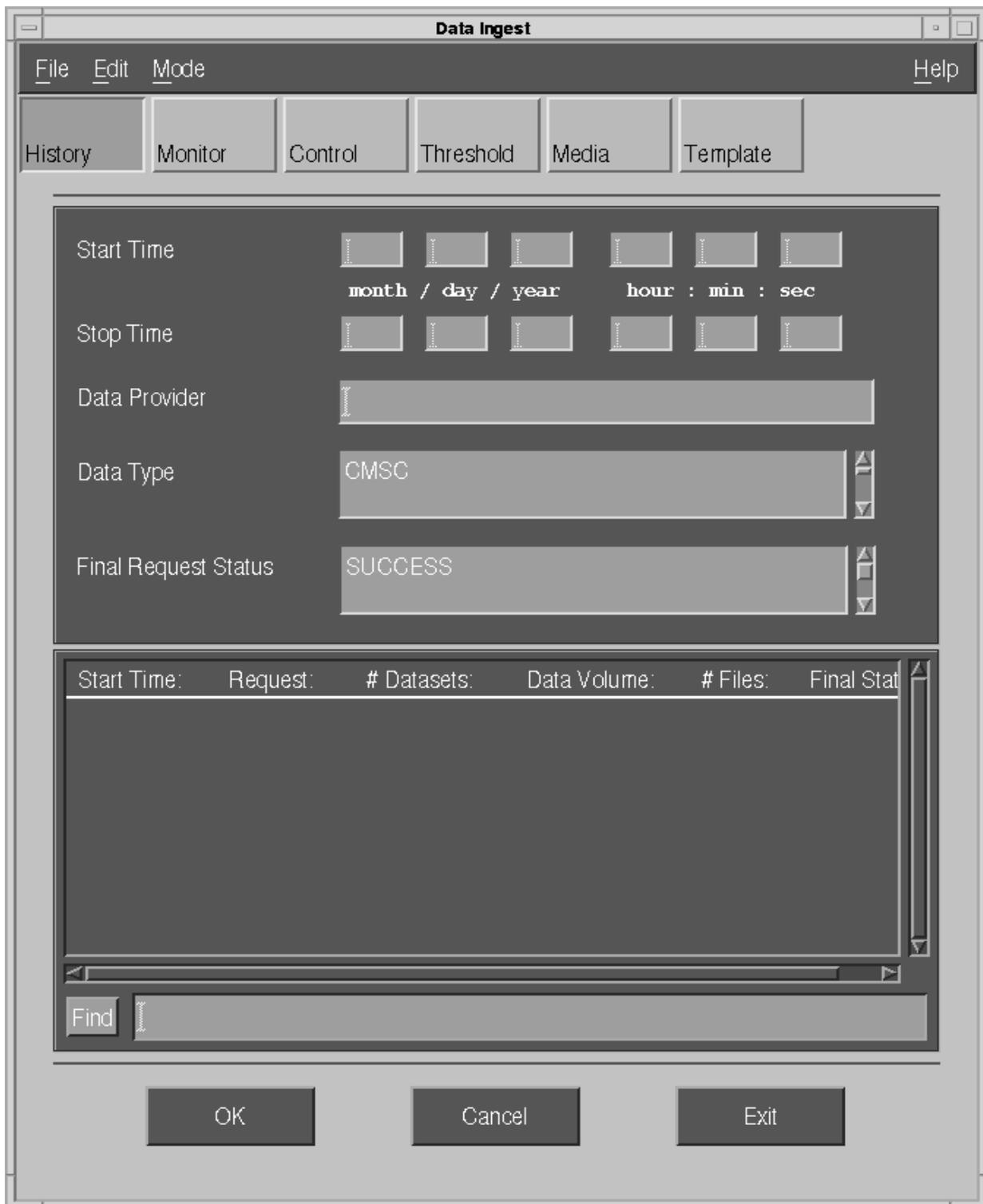
**Figure 3.9.5.6-1. Main Ingest GUI Screen**



**Figure 3.9.5.6-2. Ingest Status Monitoring GUI Screen**



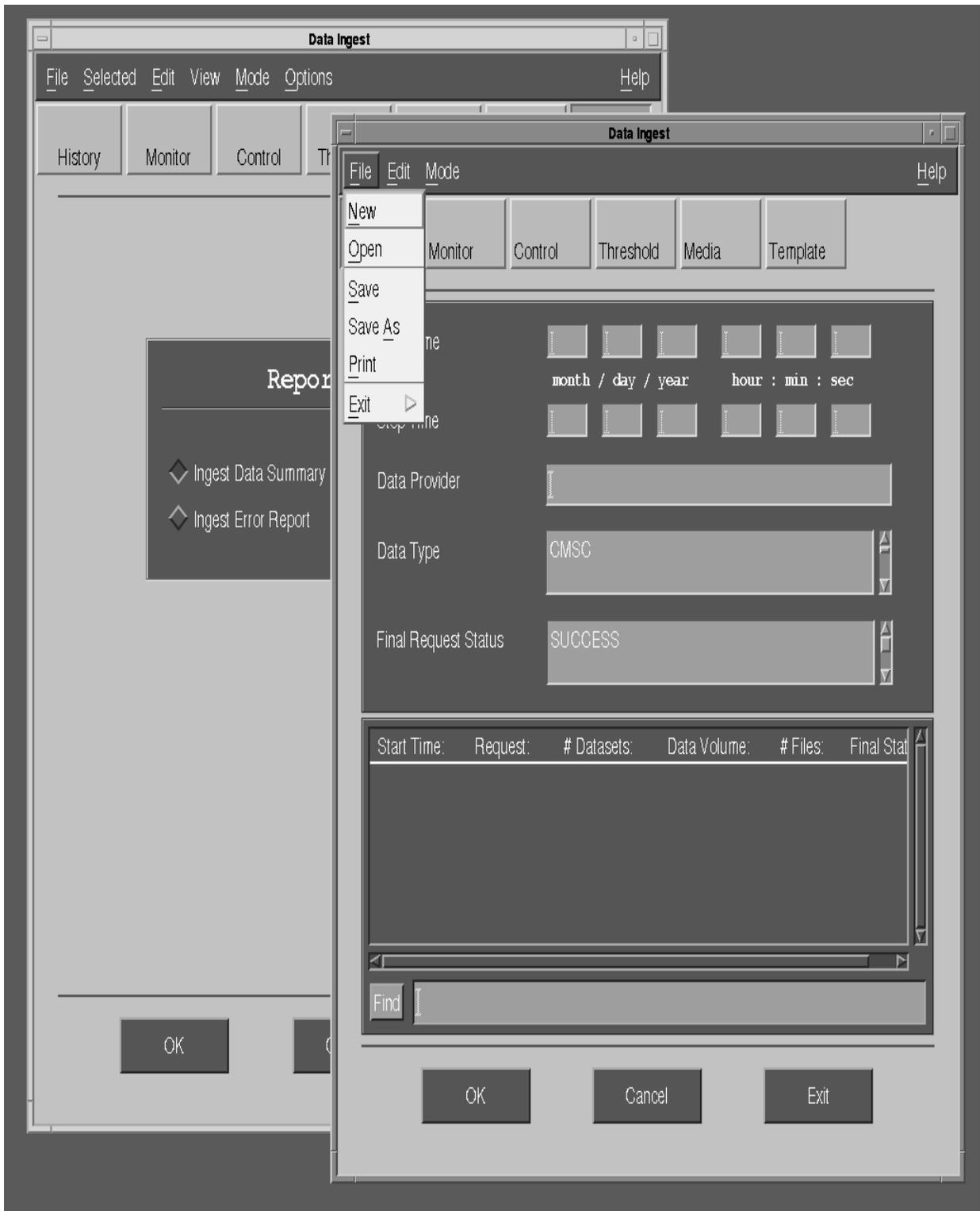
**Figure 3.9.5.6-3. Ingest Request Cancellation GUI Screen**



**Figure 3.9.5.6-4. Ingest History Log GUI Screen**



**Figure 3.9.5.6-5. Ingest Summary Report GUI Screen**



**Figure 3.9.5.6-6. Ingest Summary Report GUI Selection Screen**